

Progress in the Application of DNA Microarrays

Edward K. Lobenhofer, Pierre R. Bushel, Cynthia A. Afshari, and Hisham K. Hamadeh

Laboratory of Molecular Carcinogenesis, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA

Microarray technology has been applied to a variety of different fields to address fundamental research questions. The use of microarrays, or DNA chips, to study the gene expression profiles of biologic samples began in 1995. Since that time, the fundamental concepts behind the chip, the technology required for making and using these chips, and the multitude of statistical tools for analyzing the data have been extensively reviewed. For this reason, the focus of this review will be not on the technology itself but on the application of microarrays as a research tool and the future challenges of the field. **Key words:** DNA chips, gene expression profiling, microarray, toxicogenomics. *Environ Health Perspect* 109:881–891 (2001). [Online 16 August 2001] <http://ehpnet1.niehs.nih.gov/docs/2001/109p881-891lobenhofer/abstract.html>

The Herculean effort to sequence every nucleotide in a genome has generated complete sequences for many of the common experimental organisms, including *Saccharomyces cerevisiae*, *Drosophila melanogaster*, and *Caenorhabditis elegans*. Significant progress has also been made sequencing entire genomes for more complex species, such as mice and humans. The analysis and application of this information have the potential to advance rapidly our understanding of the molecular mechanisms underlying normal cellular processes as well as the molecular basis for disease. A limitation of this information is that complete sequence data merely enable the identification of all theoretical genes within a genome. Sequence information alone does not indicate either the expression patterns for a given gene or the cell type in which it is expressed.

The functional expression of genes may be assessed by measuring the levels of mRNA and their corresponding proteins. Previously, one could assess the expression of mRNA for known genes or unidentified genes or expressed sequence tags (ESTs) using somewhat labor-intensive, low-throughput techniques, including Northern blot analysis and RNase protection assays. The advent of newer technologies such as quantitative reverse transcription-polymerase chain reaction (RT-PCR) and DNA chips, or microarrays, has enabled rapid and simultaneous comparison of mRNA levels for hundreds to thousands of genes in virtually any biologic sample.

The number of publications employing microarrays has undergone exponential growth since its inception in 1995. The original paper by Schena et al. (1) detailing this technique to monitor gene expression changes has been referenced over 700 times. Furthermore, a search of published literature using the keyword “microarray” or “DNA chip” yields over 800 articles (Figure 1). Most of these articles have focused on studies

related to cancer research. Microarrays have also been used to study questions in pathology, cell biology, pharmacology, and toxicology. The technology and methods involved in microarray experiments have been the focus of several reviews (2–9). Therefore, we will summarize briefly the principle of microarray, and then compile a summary of studies in which the application of DNA microarrays has addressed critical questions in different scientific fields as well as discuss several future challenges. Due to the vast nature of this task, this review is meant to provide a comprehensive but not inclusive summary of all expression array studies.

General Review of Microarray Technology

DNA microarray technology is revolutionary because it provides a platform to perform genome-wide expression analyses across various biologic models. The recent popularity of microarrays stems from the wide variety of research areas into which this technology can be integrated. Because physiologic responses involve complex regulatory networks that affect the levels of gene expression, the use of DNA microarrays to monitor simultaneously the response of thousands of genes allows one to observe genes that act in a coordinate fashion. DNA microarrays will help us better understand the mechanisms of action of compounds, provide insight into functions for genes with no known function, and advance many areas of biologic and biomedical research.

DNA microarrays are generated by either printing presynthesized cDNAs (500–2000 bases) or synthesizing short oligonucleotides (20–50 bases) onto glass microscope slides or membranes. cDNAs for microarrays may include fully sequenced genes of known function or collections of partially sequenced cDNA derived from expressed sequence tags (ESTs) corresponding to the messenger RNAs of unknown genes. Differential gene

expression measurements are achieved by competitive, simultaneous hybridization of reverse-transcribed cDNAs using a two-color fluorescence labeling approach or comparison of two biotin or radioactively labeled samples hybridized on different chips. One may compare, for example, two different tissues, normal versus diseased tissue or untreated versus exposed cells. Scanned images representing the two samples are then overlaid using specialized image-processing software that assigns a color intensity corresponding to the amount of fluorescence or radioactivity. The ratio of the intensity for one sample compared to the other is used as a measurement of whether a gene is significantly different (i.e., induced or repressed) in one sample from the other.

Oncology

In the past five years, microarray technology has been used as a tool in many studies in cancer research. Many of these reports are listed in Table 1, which categorizes the studies by tumor location. Microarrays have been used in cancer research to address three main objectives: determining the molecular differences between normal and malignant cells; improving the classification of tumors to increase the effectiveness of therapeutics; and identifying mutations in genes that are implicated in tumor formation or progression.

Microarray technology has revolutionized the manner in which cancer researchers identify genes that are differentially expressed during tumor progression. Using breast tissue that was obtained immediately after a modified radical mastectomy, Sgroi et al. (10) examined differences in gene expression in the tumor and normal tissue from a single patient. The tissue was frozen, sectioned, and stained using hematoxylin and eosin to distinguish normal breast epithelial cells from invasive carcinoma cells as well as metastatic cells (from an axillary lymph node). Each cell population (normal, primary tumor, metastatic tumor) was isolated using laser capture

Address correspondence to C.A. Afshari, National Institute of Environmental Health Sciences, P.O. Box 12233, Research Triangle Park, NC 27709 USA. Telephone: (919) 541-1310. Fax: (919) 541-0146. E-mail: afshari@niehs.nih.gov

We thank L.T. Bennett, C.J. Tucker, R. Amin, J.B. Collins, G.A. Boorman, and T.R. Devereux for critically reviewing this manuscript as well as providing insightful comments.

Received 4 January 2001; accepted 12 March 2001.

microscopy to avoid cross-contamination, inclusion of adipose tissue, stromal components, or lymphocytes that may have been associated with an immune response to the disease. Isolation of different cell populations from a single patient meant that any observed gene expression changes could potentially be attributed directly to disease progression, a task that would be more difficult if the study had been performed using tissue from multiple women. Sufficient numbers of each cell type were harvested using laser capture microscopy to yield enough RNA for microarray analysis without using RNA amplification procedures (11,12). (Appropriate cell numbers will vary depending upon the cell type and tissue of origin, and the amount of RNA depends on the microarray platform and the labeling protocol.) RNA amplification should be used with caution to avoid the risk of misrepresentational gene expression biases (12–14).

This study (10) identified genes that were differentially expressed in invasive or metastatic breast cancer cells compared with levels present in normal breast epithelial cells. To increase the statistical confidence in the data, the analysis included data from replicate hybridizations. Many of the differentially expressed genes were verified by quantitative RT-PCR and immunohistochemical assays. A significant proportion (41%, 83 of 202) of the differentially expressed genes were unknown; however, many of the known genes were previously implicated in tumor biology or specifically in breast cancer.

For example, the increased expression of tissue factor (also known as thromboplastin or coagulation factor III) has been described

previously during angiogenesis and metastasis [reviewed by Ruf and Mueller (15)]. In a more recent study (16), increased levels of tissue factor protein have been detailed in the tumor, stroma, and plasma of breast cancer patients. The microarray data demonstrated a 5-fold increase in tissue factor transcripts in primary tumor cells relative to normal breast epithelial cells. This finding corroborated the association of increased levels of tissue factor with tumor progression, and also indicated that tissue factor expression is under transcriptional control. Cumulatively, this study demonstrated that the use of carefully controlled experimental parameters (unamplified RNA from distinct cell populations obtained from one individual patient) in the preparation of samples for microarray analysis can yield a descriptive examination of the transcriptional changes associated with tumor progression.

Microarrays have also been used to study tumor progression in cell cultures. It has been reported (17) that the human melanoma cell line UACC-903 can be reverted to a normal growth rate after introduction of a normal human chromosome 6 [UACC-903(+6)]. In a more recent study (18), fluorescently labeled cDNA probes were generated from mRNA extracted from UACC-903 or UACC-903(+6) cells and hybridized to a DNA chip to identify changes in gene expression associated with the reduced rate of proliferation and tumor suppressed phenotype. The criteria used to determine whether a gene was differentially expressed identified 63 upregulated genes (7.3% of the 870 genes present on the chip) on introduction of normal chromosome 6, whereas 15 (1.7%) genes

were significantly downregulated. The levels of 15 of these differentially expressed genes were verified independently by Northern blot analysis. One of the validated genes demonstrating increased expression in the UACC-903(+6) line was *p21/WAF1*, a mediator of p53-induced growth suppression, which correlated with the reduced growth rate observed in transfected cells. This report exemplifies the quality and usefulness of microarray data and provides further validation of the potential utility of *in vitro* models of cancer.

Microarray technology can facilitate accurate classifications of cancer. Historically tumors have been classified by pathologic examination supplemented by special stains and antibodies. Similar to the way in which DNA sequencing is revolutionizing the field of taxonomy, gene expression profiling is increasing the number of markers useful in tumor classification. For example, because of standard immunohistochemical techniques that showed positive staining for neural-specific markers, small cell lung cancers were believed to originate from neuroendocrine tissue (19). Pulmonary carcinoid tumors, which are benign lung tumors, also stain positively for neuroendocrine markers and are therefore hypothesized to arise from neuroendocrine cells. Using high-density cDNA arrays, Anbazhagan et al. (20) supported this theory by demonstrating that pulmonary carcinoid tumors exhibited a gene expression profile similar to that of two neural-derived tumors, oligodendroglioma and high-grade astrocytoma. Surprisingly, the expression profile of small cell lung cancer was more similar to that of bronchial epithelial cells than to those of carcinoid tumors, oligodendrogliomas, or high-grade astrocytomas. This finding suggests that small cell lung cancer may actually be an epithelial tumor rather than of neuroendocrine origin. In studies of melanomas, diffuse large B-cell lymphomas,

Table 1. Use of microarray analysis in cancer research.

Location of primary tumor/ tumor type	Reference
Breast	(10,26,28,29,31,81–96)
Ovary/cervix/endometrium	(32,97–100)
Prostate	(101–109)
Colon	(29,110–115)
Pancreas	(116,117)
Kidney	(118)
Blood/lymph	(21,23,119–129)
Liver	(130,131)
Bladder	(132,133)
Head/neck/oral	(134–137)
Brain	(138–143)
Lung	(20,29,144–149)
Skin	(18,22,150–152)
Rhabdomyosarcoma	(153–156)
Ewing's sarcoma	(157)
Multiple sites/tumor types	(63,64,158–161)

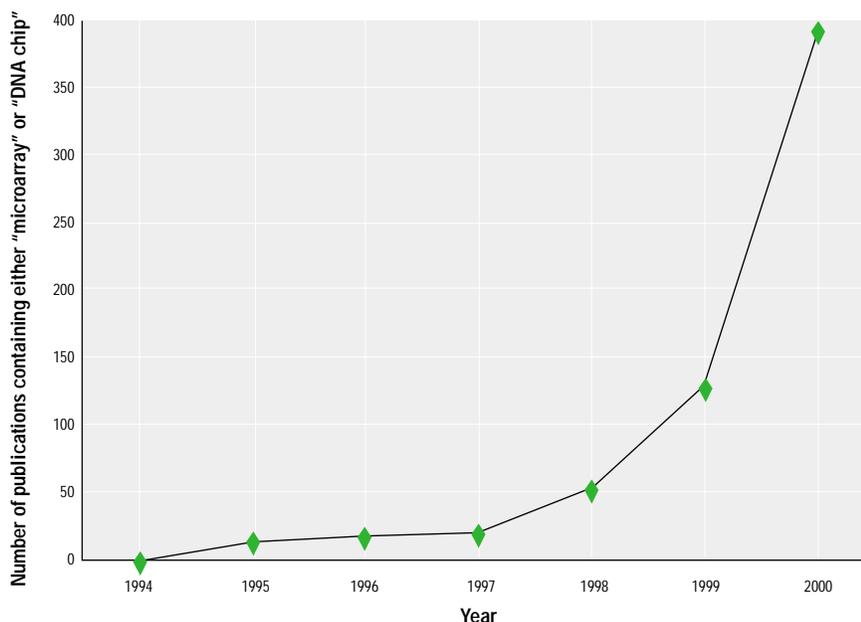


Figure 1. Exponential growth of microarray technology use in scientific reports. Articles published per year containing the keyword "microarray" or "DNA chip" were identified by a Medline search.

and leukemias, previously uncharacterized tumor subtypes were identified based on differential gene expression patterns observed using microarrays (21–23). Ultimately, more accurate classification of cancerous lesions based on gene expression profiles may improve the therapeutic strategies employed to treat patients by allowing selection of chemotherapeutic agents that target tumor subtype-specific molecules.

Breast cancers are currently classified based on several parameters, including tumor size and stage as well as the level of estrogen receptor (ER) expression (24,25). Using several breast cancer cell lines as well as primary tumor samples, Martin et al. (26) identified clusters on a hierarchical tree (groups of genes that demonstrate similar expression patterns) that are associated with each of these clinical features. Grouping—statistical clustering of tumors according to their respective gene expression profiles—enabled discrimination of tumors based on ER status. However, several tumors were not grouped in accordance with the original immunohistochemical data. Not a shortcoming of the gene expression profiling procedure, this discrepancy represents a strength of this technique by providing additional insight into tumor biology. Some tumors are classified as ER-positive based on immunohistochemical analysis, but actually express a nonfunctional protein and therefore should be typed ER-negative. This would indicate that tumors expressing ER (determined by classic assays) but not grouped with other ER-positive tumors by clustering of gene expression data might represent this unique group of tumors. The accurate classification of these tumors using a combination of classic pathologic as well as gene expression techniques would enable physicians to predict whether these tumors would be refractive to hormonal therapy, and allow the consideration of alternative treatment regimes.

Advances in cancer research can be facilitated by studies that use microarrays to explore the function of genes implicated in tumor progression or genetic susceptibility. For example, women with a germline mutation in *BRCA1* are predisposed to breast and ovarian cancer (27). *BRCA1* has been implicated in the response to DNA damage, homologous recombination, and transcription. To isolate target genes of *BRCA1* action, two different laboratories completed high-density array hybridizations using RNA from cells ectopically expressing *BRCA1* (28,29). Consistent with the proposed action of *BRCA1* in DNA repair, both groups identified genes implicated in the DNA damage response (such as *GADD45*) as being significantly upregulated in cells overexpressing *BRCA1*. Further exploration

into the molecular mechanisms of *BRCA1* action and the effect of gene mutations on its activity will undoubtedly provide more insight into the basis for genetic predisposition to cancer.

To detect gene mutations, a variation on array-based technology has been developed. *N*-mer arrays contain oligonucleotides of specific length, which correspond to a portion of the coding region of the gene of interest. Wild-type sequences as well as single base substitutions at every position are synthesized on a chip. After hybridization, the intensity of each oligonucleotide is assessed; the oligonucleotide with the highest signal corresponds to the actual sequence for the gene being analyzed (30). This technique has identified single base insertions, deletions, and substitutions in three different tumor suppressor genes, *BRCA1*, *BRCA2*, and *p53* (31,32). This modified application of microarray technology has the promise to deliver high-throughput data with the same accuracy as more popular, gel-based methods of mutation screening while being less time- and labor-intensive (32).

Infectious Disease

The underlying basis of infectious disease stems from the complex interaction between a pathogen and a host. Understanding the molecular details of this interaction will aid in the identification of virulence-associated microbial genes and host-defense strategies. This would enable characterization of the genes involved in pathogenesis and the regulatory mechanisms controlling these genes. Ultimately, this knowledge could guide the generation of new treatment regimes to combat infections. Microarray technology promises to expedite microbial research by aiding in annotation of the microbial genome, examination of microbial physiology, and identification of candidate virulence factors.

Most of the time, the function of bacterial genes is inferred by finding homologs that have been previously characterized in other organisms. The study of genomewide expression patterns will greatly enhance the annotation of gene function: Similar expression may imply related function. Thus, when a gene with an unknown function is clustered with known genes based on similar expression patterns, it is implied that the function of the unknown gene is related to that of other genes in that cluster. This hypothesis was first tested in a study examining the gene expression profile of *Saccharomyces cerevisiae*, and it demonstrated the coregulated expression of genes that performed similar cellular functions (33). Theoretically, expression profiling will enable the function of unknown genes to be extrapolated, even when sequence similarity

has not been established with characterized genes. This is especially true for pathogens because the whole genome has been sequenced for a significant number of these organisms using rapid techniques such as the “shotgun” approach (34–36).

A series of studies demonstrated that the evaluation of mRNA expression profiles by DNA microarray analysis is a powerful approach for characterizing and understanding host–pathogen interactions (37–43). The ability to identify virulence factors using DNA microarrays has been shown by comparing gene expression of pathogen-infected and -uninfected cells. Cohen et al. (39) observed changes in the gene expression profile of the human promyelocytic THP1 cell line after infection with *Listeria monocytogenes*, a pathogenic intracellular organism responsible for meningitis, meningoencephalitis, and in some cases gastroenteritis. The underlying premise of their study design was that this gram-positive bacterium is able to penetrate and grow in both phagocytes and nonphagocytic target cells. The infection stimulated transient activation of signal transduction pathways in the host cell, leading to modulation of gene expression. The authors' findings were consistent with previously published data, and also identified several novel genes associated with the infection. *L. monocytogenes*-regulated genes belonged to different functional categories, including inflammation, chemotaxis, transcription, apoptosis, transduction, metabolism, and cell cycle. This information will enhance the understanding of the molecular physiology of *L. monocytogenes* and guide the development of new therapeutic drugs.

Cellular Biology

Microarray technology has influenced cell biology by enabling characterization of different cellular events at the molecular level (Table 2). This information has increased the understanding of major biologic events such as organ development, homeostasis, as well as inter- and intraorgan communication.

Studies using microarrays have not only enhanced the understanding of known cellular processes, but also aided the discovery of previously unknown properties or functions of different cells. Iyer et al. (44) demonstrated this through analysis of the response of human fibroblast cells to serum. Genes were clustered into groups on the basis of their temporal patterns of expression. Many features of the transcriptional program appeared to be related to the physiology of wound repair, suggesting that fibroblasts may play a larger role in this complex, multicellular response than had previously been assumed.

When studying human responses, the goal is to use human-derived tissue that is

representational of the organ or system being studied. However, several gene expression studies have identified differences between widely used cell culture models and the tissue of origin. Detailed studies regarding the function of normal cells will help identify limitations of cell models. For example, Walker and Rigley (45) examined gene expression profiles in human peripheral blood mononuclear cells (PBMCs) stimulated with phytohemagglutinin. They found 104 genes to be differentially expressed in response to phytohemagglutinin stimulation. Clustering analysis grouped these genes into major functional categories including detoxification, intracellular signaling, vesicle trafficking, inflammation, and house keeping. Due to the vast number of studies using PBMCs, it is important that changes in their gene expression profile resulting from different treatment and culturing conditions be studied and characterized. This information will aid greatly in interpreting data from studies using these cells and identify changes in gene expression resulting from variation in the model.

Additional complications of *in vitro* studies are that cells can differentiate, grow old, or become transformed during experimentation through treatment or growth conditions. Cell differentiation during an *in vitro* study introduces potential biological

variability into a data set. Changes in gene expression during differentiation have been studied by comparison of differentiated 3T3-L1 adipocytes and nondifferentiated 3T3-L1 preadipocytes (46). A vast difference in expression patterns was discovered, highlighting the importance of identifying experimental parameters that can confound experimental results.

DNA microarrays have also been used to better understand mechanisms associated with aging by identifying specific genes associated with age-related processes. One study (47) evaluated gene expression alterations in *in vitro*-aged dermal fibroblast cell populations following immortalization with the telomerase catalytic subunit (hTERT). Telomeres are specialized DNA repeats that stabilize the ends of chromosomes (48,49). As somatic cells divide, their telomeres gradually become shorter until a critical length is reached, after which somatic cells enter a nondividing state termed senescence (50). Microarray analysis showed that senescent cells express reduced levels of collagen I and III, as well as increases in genes associated with the destruction of dermal matrix and inflammatory processes. The study also demonstrated that expression of telomerase in the normal cells produced gene expression patterns very similar to early passage cells, consistent with the ability of these cells to

proliferate and perform normal biologic functions. The investigators indicated that telomerase activity not only conferred immortality in skin fibroblasts, but also reversed the loss of biologic function in senescent cells.

A separate study analyzed the gene expression profile of the aging process in the skeletal muscle of mice (51). The data revealed that aging produced a heightened stress response and lower expression of metabolic and biosynthetic genes. These findings are consistent with the senescent phenotype because decreased levels of biologic activity characterize these cells. Interestingly, these investigators also demonstrated that changes in the gene expression profile were prevented by a calorie-restrictive diet, suggesting that increasing protein turnover and decreasing macromolecular damage could extend the aging process.

Use of microarrays to examine changes in gene expression patterns can also identify regulatory genes and other essential components of different cellular processes. Monitoring RNA expression levels in organisms for which the entire genome has been sequenced (such as yeast) enables a comprehensive examination of functional genomic responses compared to other biologic models where only partial sequence is known. Lashkari et al. (52) generated the first high-density cDNA microarray of yeast open reading frames, whereas Wodicka et al. (53) developed the first high-density oligonucleotide arrays for monitoring the expression of nearly all yeast genes. The sheer volume of data generated using the *S. cerevisiae* model, including development of mutation, knock-out, and overexpression models, reduces the challenges of testing hypotheses and making associations inferred from gene expression analyses.

Eisen et al. (33) grouped genes expressed in budding yeast into functional subsets by combining gene expression profiling with clustering analysis, using an algorithm designed to arrange genes according to similarity in expression patterns. This study revealed the versatility of microarrays and highlighted the computational tools that can be applied to these data so as to derive biologic associations.

The use of genetic manipulations, in addition to genome-wide transcript profiling in yeast, revealed changes in gene expression underlying pheromone signaling, cell cycle control, and polarized morphogenesis (54). These experiments showed cross talk and overlap of multiple mitogen-activated protein kinase pathways such as filamentous growth and mating responses. More globally, this study demonstrated the potential to correlate gene expression changes with alterations in physiologic or developmental processes.

Table 2. Application of microarray technology to different scientific disciplines.

Discipline, investigation topic	Reference
Pathology	
<i>Pseudomonas aeruginosa</i>	(38)
Human papilloma virus	(42)
Human immunodeficiency virus	(39)
<i>Listeria monocytogenes</i>	(41)
Salmonella	(40)
Tuberculosis	(43)
Malarial parasite	(162)
<i>Bordetella pertussis</i>	(37)
Toxicology/pharmacology	
Carbon tetrachloride (HepG2)	(163)
Beta-naphthoflavone (mouse liver)	(164)
2,3,7,8-Tetrachlorodibenzo- <i>p</i> -dioxin (HepG2)	(165)
Copper regulation in <i>S. cerevisiae</i>	(166)
Multidrug resistance in <i>S. cerevisiae</i>	(167)
Isoniazid	(168)
Cell biology/cellular signaling	
Yeast cell cycle	(169)
Yeast adaptation to growth conditions	(170–172)
Transcription control	(173–179)
Functional diversity among kinase A subunits	(180)
Role of histones in telomeric silencing	(181)
Kinase inhibitors	(182)
Effect of clotting factor on fibroblasts	(183)
Differentiation of embryonic stem cells	(184)
Erythroid maturation	(119)
Maturation of human dendritic cells	(185)
Response of pancreatic cells to glucose	(186)
Mechanically induced genes in human muscle cells	(187)
Mid-gestation placenta and embryo associated genes	(188)
Metamorphosis	(189)
Functional diversity of T cells	(190)
Transcript profile of hematopoietic stem cells	(191)

Numerous other studies have focused on elucidating components of signaling cascades in biologic models (Table 2). Cumulatively, these studies contribute to a greater understanding of the molecular communication within cells and how those signals mediate cellular processes.

Toxicology and Pharmacology

Identification of environmental carcinogens and other hazards constitutes a major challenge in the field of toxicology. The assessment of risk associated with different chemical exposures is limited partially by difficulties relating to cross-species extrapolation and dose response as well as estimations of exposure levels. Thus, one must identify trans-species, chemical-specific biomarkers that result from quantified exposures to drugs or toxic chemicals. This would improve the extrapolation of data from model systems to potential effects in humans. Many disciplines employ *in vitro* models to study cell processes and signals; however, to extrapolate findings to humans, toxicology and pharmacology require the use of relatively large, complex organisms such as rodents, rabbits, canines, and monkeys. The process of gene discovery and deciphering these genomes is challenging, but current efforts to sequence the genomes of these organisms will greatly enhance the DNA microarray field by offering more defined sequences to monitor in the response of a biologic model to a chemical exposure.

The potential of DNA microarray technology to identify gene expression changes associated with toxic or pharmacologic processes has been the focus of several studies. For example, the biochemical and molecular mechanisms of lead neurotoxicity have not been fully elucidated; however, lead has been shown to interfere with several signal transduction pathways (55–58). Changes in gene expression in target cells is hypothesized to be a mechanism by which this chemical interferes with normal brain development (59). Hossain et al. (60) studied the mechanisms underlying lead neurotoxicity using cDNA microarray gene expression analysis and identified lead-sensitive genes in immortalized human fetal astrocytes (SV-FHA). Their findings indicated that lead induces vascular endothelial growth factor (VEGF) expression in SV-FHAs via a pathway involving protein kinase C and the AP-1 transcription factor, yet independent of hypoxia-inducible factor-1. This report highlights the potential of DNA microarrays for the discovery of novel toxicant-induced gene expression alterations and the ability to dissect the second messenger pathways and transcription factors mediating these changes. By allowing determination of the

mechanisms of toxicant action, such studies will ultimately contribute to models developed for risk assessment in humans.

Microarray experiments using pharmacologic agents have focused on identification of mechanisms for drug action as well as isolation of potential drug targets. A landmark study (61) described a method for drug target validation and identification of secondary drug target effects based on genome-wide gene expression patterns. These researchers generated mutations in yeast genes encoding putative targets of the immunosuppressant drugs cyclosporin A or FK506. Expression profiles were then identified for both mutant and wild-type cells after drug treatment to verify essential components of the drug's molecular mechanism. This research revealed pathways that FK506 affected in a calcineurin- and immunophilins-independent manner by induction of drug-dependent effects in "targetless" cells. The described method permitted direct confirmation of drug targets and recognition of drug-dependent changes in gene expression, including those mediated through pathways distinct from the drug's intended target. This parallel approach in comparison of wild-type and mutant cells may help improve the efficiency of the drug development process.

Currently, there is a limited inventory of expression profiles reflecting the response to chemical treatment in biologic models. This type of information would enhance our understanding of toxicology and pharmacology. The principal hypothesis underlying toxicogenomic studies is that chemical-specific patterns of altered gene expression can be revealed using high-density microarray analysis on the tissues from treated organisms (6). As detailed earlier, gene expression-based pattern discrimination has been effective in tumor classification (21,23,62–64); however, this problem is more complex in toxicology-based studies that are confounded by target organ, dose, and time point variables. This challenge was demonstrated in a recently published effort to distinguish classes of toxicants based solely upon cluster-type analysis of differentially expressed genes in HepG2 cells (65). Initial comparison of gene expression profiles for cytotoxic anti-inflammatory drugs and DNA-damaging agents to a database populated with gene expression profiles from 100 toxicants failed to differentiate between the two classes of compounds. The authors suggest that a lack of replicate hybridizations resulted in low reproducibility of the gene responses. To surmount this challenge, they generated a single gene-expression profile from 13 separate hybridizations using RNA isolated independently from cisplatin-treated HepG2 cultures, producing a set of genes that demonstrated altered expression

with relatively high reproducibility. This highlights the need for multiple biologic and hybridization replicates to provide statistical confidence in the gene expression profiles.

We have also verified the hypothesis that chemical-specific patterns of expression can be revealed using microarray analysis (66). We analyzed patterns of gene expression in liver tissue isolated from chemically exposed Sprague-Dawley rats using cDNA microarrays. Biologic and hybridization replicates generated statistically significant information pertaining to the expression pattern of all 1,700 genes on the DNA chip. Clustering analysis and statistical correlation revealed that gene expression profiles produced in animals treated with different agents from a common class of compounds (peroxisome proliferators) were similar. As expected, a distinct gene expression profile was produced using a compound from a completely distinct class (barbiturates). Not only did our study discriminate among different classes of toxicants on the basis of analysis of validated, statistically significant gene expression profiles, but also corroborated past findings regarding the molecular mechanisms of action for peroxisome proliferators and phenobarbital.

DNA microarrays will be instrumental in revolutionizing the fields of toxicologic and pharmacologic research. However, data generated from microarray analysis cannot be appreciated fully without computational tools that can efficiently handle and analyze large volumes of data. Thus, the field of bioinformatics has emerged and is actively addressing issues such as detecting microarray targets, calculating intensity from raw image scans, transforming data sets into more manageable forms, performing analyses that extract associations from gene expression levels, and facilitating the emergence of new testable hypotheses.

Data Analysis: Bioinformatics

The applications of database systems, computer programs, and information technology have revolutionized the way in which biologic scientists manage, analyze, and share information and data. The National Center for Biotechnology Information (www.ncbi.nlm.nih.gov) (67), a division of the National Library of Medicine at the National Institutes of Health, maintains many databases and biologic resources that contain information and data for the scientific community engaged in biomedical research. In addition, several investigative programs in computational biology have been established to leverage theoretical, analytical, and applied computer science research toward addressing complex mathematical and statistical problems inherent in modern day molecular biology. In essence,

the Human Genome Sequencing Project has ascended as a quintessential large-scale, multidisciplinary effort that has benefited immensely from the analysis, interchange,

and comparison of DNA sequencing data and results (68).

With the advent of microarray technology to monitor genomewide expression of

thousands of genes, scientists have adopted an operations research approach to analyze and manage gene expression data in a production-like setting. The essential component of

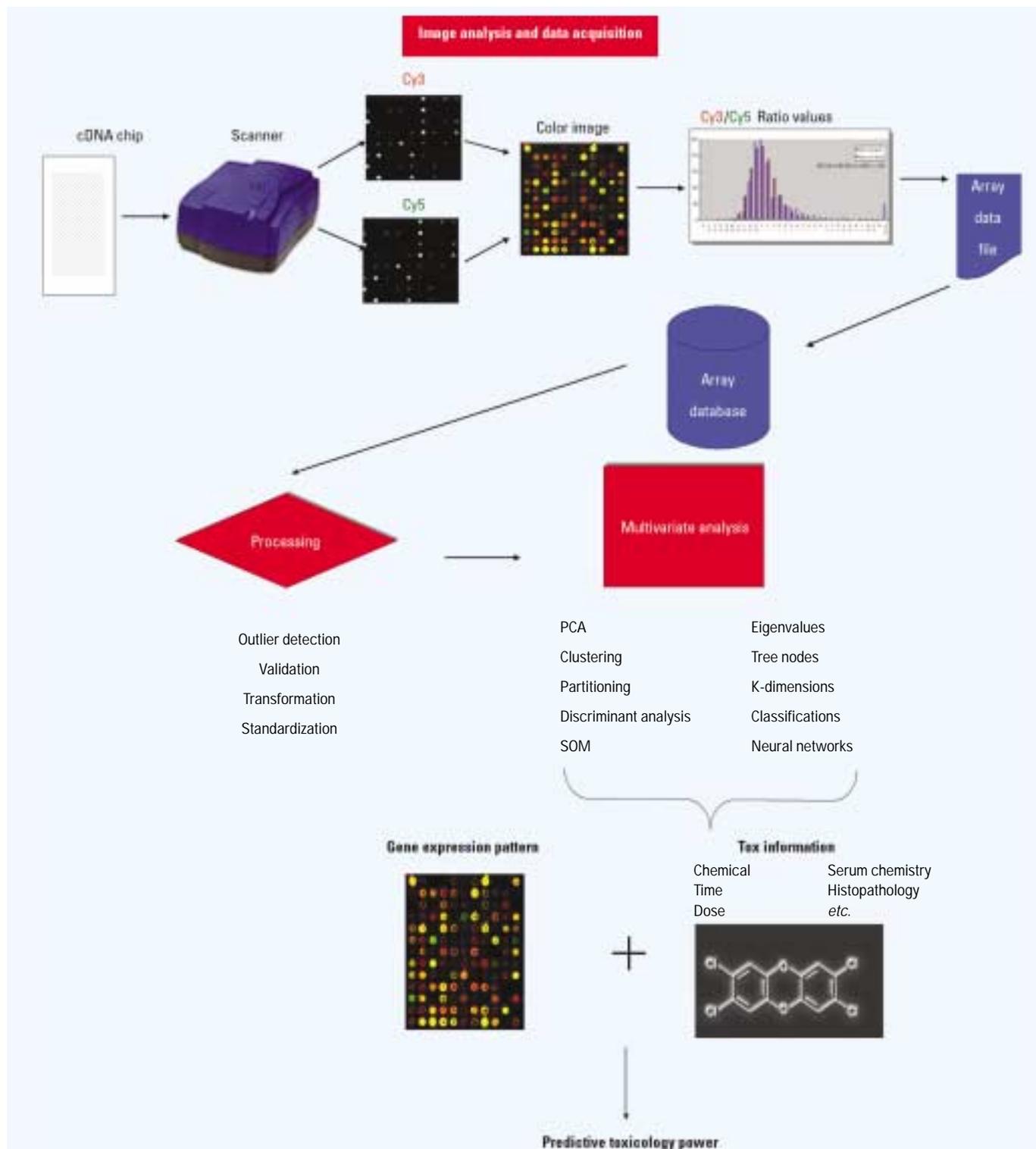


Figure 2. Example of microarray image and data analysis processes for toxicogenomic studies. Abbreviations: PCA, principal component analysis; SOM, self-organizing maps. The simplified schema illustrates the three main processes involved in microarray analysis. The cDNA chip denotes a high-density, complementary DNA glass substrate microarray. The Cy3 and Cy5 illustrations depict raw 16-bit gray scale images resulting from scanning a cDNA chip with dual lasers to excite the two cytofluorescent dyes at different wavelengths. The color image represents a standard red-green-blue (RGB), 24-bit pseudo-color composite overlay image. The Cy3/Cy5 ratio values graph illustrates the frequency distribution of the Cy3/Cy5 ratio pixel intensity values calculated for each gene target on the cDNA chip. Data then are usually transformed into a numeric data file for further processing steps as depicted.

microarray data analysis is capturing meaningful information regarding gene regulation. These data ultimately will be used to gather heuristic knowledge about a biologic system, toxic agent, or environmental condition affecting human health and disease states (5,6,9,69). The microarray data analysis schema comprises three main processes: image analysis and data acquisition, data processing, and multivariate analysis (Figure 2).

The initial stage of microarray gene expression analysis is image analysis and data acquisition. Typically a microarray is scanned in a chip reader using lasers to excite two different cytofluorescent dyes that have been incorporated into the DNA; simultaneously the chip reader collects the separate emission signals with dual photomultiplier tubes (PMT). The output current of each PMT is then digitized to form a raw 16-bit gray scale image (8,70). The most current scanning technology performs imaging in the range of 5–10 mm resolution; however, new developments in scanning technology should improve imaging resolution and ultimately enhance the signal-to-noise and dynamic range of detection.

Changes in gene expression are quantified from raw images using image analysis algorithms and data acquisition procedures to detect target regions, compute feature pixel intensity values, subtract background intensity, and compare the pixel intensity values of DNA from a treated sample to that of a control or untreated sample. Normalization or calibration steps are regularly performed in microarray data analysis to accurately adjust and correct the variability in intensity values extracted from two separate samples according to a standard set of control genes or against all the targets represented on the DNA chip. A final mean ratio pixel intensity value is calculated for each gene target and processed into a standard 24-bit pseudocolor composite image for visualization of the gene expression profile.

Given the high “gene feature” density of microarray chips, one may store pixel intensity values, calculated ratio, and statistics for each gene as records in a tab-delimited text file for subsequent data management and analysis. For instance, the microarray database ArrayDB was designed as a relational database management system to store, analyze, and associate gene expression data with information from remote biologic resources (71). More recently, a comprehensive database has been developed to correlate gene expression patterns with information about pharmacologic compounds (63).

The next stage in microarray data analysis is data processing. Inherent in microarray gene expression data is the nonlinear nature of the data. Low-quality or unreliable data

points originate from fluorescent signals below threshold detection levels or anomalies on the microarray chip, or are the result of missing DNA features. Typical microarray data analysis software includes functions to log-transform gene expression data, which essentially stabilizes the variance in the data. In addition, cutoff values may be applied to the pixel size of DNA features as well as to the intensity range of the values to refine the data set for further analysis. These processing features filter microarray gene expression data and generally improve the reliability of results.

Determination of differentially expressed genes is the staple of this processing stage, which adds another level of complexity to the analysis of microarray gene expression data, but promises to play a pivotal role in fundamental microarray gene expression pattern recognition exercises. At the onset of the microarray era, simplistic statistical methods were used to facilitate the interpretation of complex and large-scale microarray gene expression results. For example, a set number of standard deviations above and below the mean intensity values typically identified transcriptional changes in biologically relevant genes in most published microarray-related studies. In newer, more sophisticated methods, computed confidence intervals are used to determine significantly changed genes from a ratio distribution of gene expression data (72). In this application, differentially expressed genes are identified on the basis of confidence intervals determined from a probability distribution of the ratios of intensity values, with significantly changed genes being those that fall outside of the confidence limits. Ratio values close to 1.0 represent genes that are expressed similarly in the two samples being compared. More importantly, differentially expressed genes can be validated by comparing replicate measurements of microarrays and performing subsequent biologic assays to confirm the biologic significance of altered gene expression (73,74). Using this method, the probability that a “validated” gene occurred by chance at a given confidence level can be determined using a binomial distribution model (75).

At this time, these methods do not control for false positives and false negatives, low fold changes, or variation of data across independent replicate experiments, nor do they account for genes exhibiting a bias when detected from microarray fluor-reverse experiments. Fortunately, surveying changes in microarray gene expression data has recently attracted significant attention and gained interest from practitioners in statistics, mathematics, and applied physics communities as well as expert bioinformatics companies committed to applying rigorous

and innovative computational error models to measurements of microarray gene expression data. For instance, pioneering analysis of microarray gene expression data is currently underway using various mixed linear models and general analysis of variance models to assess gene statistical significance from microarray gene expression data (76,77). Similarly, a combination of robust additive and multiplicative error models is used to account for the uncertainty and variation in measuring gene expression data as well as assign confidence and probability values, weight, and error bars to define statistical significance of individual genes (78). Though these new and practical approaches emphatically meet the computational needs of the microarray community, it remains to be seen how mainstream and conventional these procedures will be in the next generation of commonplace microarray gene expression analysis tools.

The final stage in microarray gene expression analysis is multivariate analysis. In this step, higher-order computational tools and procedures assist with interpretation and visualization of complex multivariate microarray gene expression data. Although there is no definitive rule about which form of multivariate analysis to use on microarray gene expression data, it is critical to understand fully the implicit characteristics of each algorithm so as to exploit inherent advantages of each to gain the most intriguing insight and interpretation of analyzed results.

For example, hierarchical clustering methods are essentially iterative processes to organize objects into groups with similar attributes based on resemblance, proximity, and similarity or dissimilarity measurements. The four basic steps in conducting a cluster analysis on microarray gene expression data are data collection and selection of variables (genes and experiments) for analysis; generation of a resemblance matrix using a particular mathematical formula to measure the extent of similarity; selection and execution of an iterative clustering method to produce a phylogenetic tree diagram (dendrogram) with the branches representing the amount of similarity between clusters; and determination of where to “cut” the tree into a select number of nodes. Recently, robust computational tools and an interactive graphical interface have been developed to facilitate the compilation and visualization of clustered microarray gene expression data (33). Although cluster analysis is a classic and rather simple partitioning technique, its application is relatively new to microarray data analysis and is not supported by any comprehensive body of statistical literature.

Self-organizing maps (SOMs) are another useful algorithm used to interpret and visualize

large high-dimensional data sets in two-dimensional space. Basically, SOMs are an array of interconnected cells that become refined to various input signal patterns or classes of patterns in an orderly manner. These artificial neural networks undergo a competitive and unsupervised iterative learning process. When microarray gene expression data are used to train SOMs, structure is imposed on the data, with neighboring nodes defining related clusters. In essence, the SOM constitutes a new paradigm in artificial intelligence and cognitive learning and has recently been implemented into a computer program as a genetic neural network model for interpreting and displaying various patterns of gene expression data (79).

Finally, discriminant analysis is commonly used to determine which variables best distinguish between two or more groups. The general concept underlying the use of discriminant analysis for microarray data is to determine whether classes of gene expression measurements differ significantly with regard to distinct and characteristic gene expression profiles. Microarray data are used to assess and rank the importance of particular genes for discrete sample classification. When identified, hallmark gene markers are used to predict the classifications of test microarray treatments as well as apply robust pattern recognition procedures to large volumes of complex gene expression data. Computationally, discriminant analysis is similar to analysis of variance; however, a novel approach that combines a genetic algorithm and the k-nearest neighbor method has been described to positively identify genes that discriminate between classes of tumor cells (80) and toxicant treatments (66).

DNA microarrays have revolutionized the basic approach to research. This technique has generated collaborations among various scientific groups (pathologists, molecular and cellular biologists, toxicologists, and the like). Additionally, microarrays have brought together experts in engineering, bioinformatics, and statistics. The large and complex nature of these studies will facilitate the partnering of different research institutions in both the public and private sectors.

The popularity of DNA microarrays is derived from the promise that this technology will rapidly advance our understanding fundamental biologic questions. Microarrays have the potential to markedly increase our understanding of not only the process of disease but also the interactions between biologic organisms and their environment.

REFERENCES AND NOTES

- Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270:467-470 (1995).
- Brazma A, Vilo J. Gene expression data analysis. *FEBS Lett* 480:17-24 (2000).
- Brady G. Expression profiling of single mammalian cells—small is beautiful. *Yeast* 17:211-217 (2000).
- Celis JE, Kruhoffer M, Gromova I, Frederiksen C, Ostergaard M, Thykjaer T, Gromov P, Yu J, Palsdottir H, Magnusson N, et al. Gene expression profiling: monitoring transcription and translation products using DNA microarrays and proteomics. *FEBS Lett* 480:2-16 (2000).
- Pennie WD. Use of cDNA microarrays to probe and understand the toxicological consequences of altered gene expression. *Toxicol Lett* 112-113:473-477 (2000).
- Nuwaisir EF, Bittner M, Trent J, Barrett JC, Afshari CA. Microarrays and toxicology: the advent of toxicogenomics. *Mol Carcinog* 24:153-159 (1999).
- Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, Hughes JE, Snesrud E, Lee N, Quackenbush J. A concise guide to cDNA microarray analysis. *Biotechniques* 29:548-560 (2000).
- Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM. Expression profiling using cDNA microarrays. *Nat Genet* 21:10-14 (1999).
- Hamadeh HK, Bushel P, Nuwaisir EF, Paules R, Barrett JC, Afshari CA. cDNA microarray technology: merging toxicology and genomics. In: *Comments in Toxicology* (Pennie WD, Kimber I, eds). Boston:Gillette Co., in press.
- Sgroi DC, Teng S, Robinson G, LeVangie R, Hudson JR Jr, Elkahoulou AG. *In vivo* gene expression profile analysis of human breast cancer progression. *Cancer Res* 59:5656-5661 (1999).
- Eberwine J. Amplification of mRNA populations using aRNA generated from immobilized oligo(dT)-T7 primed cDNA. *Biotechniques* 20:584-591 (1996).
- Wang E, Miller LD, Ohnmacht GA, Liu ET, Marincola FM. High-fidelity mRNA amplification for gene profiling. *Nat Biotechnol* 18:457-459 (2000).
- Baugh LR, Hill AA, Brown EL, Hunter CP. Quantitative analysis of mRNA amplification by *in vitro* transcription. *Nucleic Acids Res* 29(5):E29 (2001).
- Madison RD, Robinson GA. lambdaRNA internal standards quantify sensitivity and amplification efficiency of mammalian gene expression profiling. *Biotechniques* 25:504-512 (1998).
- Ruf W, Mueller BM. Tissue factor in cancer angiogenesis and metastasis. *Curr Opin Hematol* 3:379-384 (1996).
- Ueno T, Toi M, Koike M, Nakamura S, Tominaga T. Tissue factor expression in breast cancer tissues: its correlation with prognosis and plasma concentration. *Br J Cancer* 83:164-170 (2000).
- Trent JM, Stanbridge EJ, McBride HL, Meese EU, Casey G, Araujo DE, Witkowski CM, Nagle RB. Tumorigenicity in human melanoma cell lines controlled by introduction of human chromosome 6. *Science* 247:568-571 (1990).
- DeRisi J, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, Chen Y, Su YA, Trent JM. Use of a cDNA microarray to analyze gene expression patterns in human cancer. *Nat Genet* 14:457-460 (1996).
- Guinee DG Jr, Fishback NF, Moss MN, Abbondanzo SL, Travis WD. The spectrum of immunohistochemical staining of small-cell lung carcinoma in specimens from transbronchial and open-lung biopsies. *Am J Clin Pathol* 102:406-414 (1994).
- Anbazhagan R, Tihan T, Bornman DM, Johnston JC, Saltz JH, Weigering A, Piantadosi S, Gabrielson E. Classification of small cell lung cancer and pulmonary carcinoid by gene expression profiles. *Cancer Res* 59:5119-5222 (1999).
- Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403:503-511 (2000).
- Bittner M, Meltzer P, Chen Y, Jiang Y, Sefter E, Hendrix M, Radmacher M, Simon R, Yakhini Z, Ben-Dor A, et al. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature* 406:536-540 (2000).
- Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286:531-537 (1999).
- Winchester DP, Cox JD. Standards for diagnosis and management of invasive breast carcinoma. *CA-Cancer J Clin* 48:83-107 (1998).
- Donegan WL. Tumor-related prognostic factors for breast cancer. *CA Cancer J Clin* 47:28-51 (1997).
- Martin KJ, Kritzman BM, Price LM, Koh B, Kwan CP, Zhang X, Mackay A, O'Hare MJ, Kaelin CM, Mutter GL, et al. Linking gene expression patterns to therapeutic groups in breast cancer. *Cancer Res* 60:2232-2238 (2000).
- Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, et al. A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science* 266:66-71 (1994).
- Harkin DP, Bean JM, Miklos D, Song YH, Truong VB, Englert C, Christians FC, Ellisen LW, Maheswaran S, Oliner JD, et al. Induction of GADD45 and JNK/SAPK-dependent apoptosis following inducible expression of BRCA1. *Cell* 97:575-586 (1999).
- MacLachlan TK, Somasundaram K, Sgagias M, Shifman Y, Muschel RJ, Cowan KH, El-Deiry WS. BRCA1 effects on the cell cycle and the DNA damage response are linked to altered gene expression. *J Biol Chem* 275:2777-2785 (2000).
- Gunderson KL, Huang XC, Morris MS, Lipshutz RJ, Lockhart DJ, Chee MS. Mutation detection by ligation to complete n-mer DNA arrays. *Genome Res* 8:1142-1153 (1998).
- Favis R, Day JP, Gerry NP, Phelan C, Narod S, Barany F. Universal DNA array detection of small insertions and deletions in BRCA1 and BRCA2. *Nat Biotechnol* 18:561-564 (2000).
- Wen WH, Bernstein L, Lescalleet J, Beazer-Barclay Y, Sullivan-Halley J, White M, Press MF. Comparison of TP53 mutations identified by oligonucleotide microarray and conventional DNA sequence analysis. *Cancer Res* 60:2716-2722 (2000).
- Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 95:14863-14868 (1998).
- Bowman S, Horrocks P. Assessing the impact of *Plasmodium falciparum* genome sequencing. *Microbes Infect* 2:1479-1487 (2000).
- Duffieux F, Peyret P, Roe BA, Vivares CP. First report on the systematic sequencing of the small genome of *Encephalitozoon cuniculi* (Protozoa, Microspora): gene organization of a 4.3 kbp region on chromosome I. *Microb Comp Genomics* 3:1-11 (1998).
- Gardner MJ, Tettelin H, Carucci DJ, Cummings LM, Smith HO, Fraser CM, Venter JC, Hoffman SL. The malaria genome sequencing project: complete sequence of *Plasmodium falciparum* chromosome 2. *Parassitologia* 41:69-75 (1999).
- Ichikawa JK, Norris A, Bangera MG, Geiss GK, van 't Wout AB, Bumgarner RE, Lory S. Interaction of *Pseudomonas aeruginosa* with epithelial cells: identification of differentially regulated genes by expression microarray analysis of human cDNAs. *Proc Natl Acad Sci USA* 97:9659-9664 (2000).
- Chang YE, Laimins LA. Microarray analysis identifies interferon-inducible genes and Stat-1 as major transcriptional targets of human papillomavirus type 31. *J Virol* 74:4174-4182 (2000).
- Cohen P, Bouaboula M, Bellis M, Baron V, Jbilou O, Pointot-Chazel C, Gallegue S, Hadibi EH, Casellas P. Monitoring cellular responses to *Listeria monocytogenes* with oligonucleotide arrays. *J Biol Chem* 275:11181-11190 (2000).
- Behr MA, Wilson MA, Gill WP, Salamon H, Schoolnik GK, Rane S, Small PM. Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* 284:1520-1523 (1999).
- Eckmann L, Smith JR, Housley MP, Dwinell MB, Kagnoff MF. Analysis by high density cDNA arrays of altered gene expression in human intestinal epithelial cells in response to infection with the invasive enteric bacteria *Salmonella*. *J Biol Chem* 275:14084-14094 (2000).
- Geiss GK, Bumgarner RE, An MC, Agy MB, van 't Wout AB, Hammersmark E, Carter VS, Upchurch D, Mullins JI, Katze MG. Large-scale monitoring of host cell gene expression during HIV-1 infection using cDNA microarrays. *Virology* 266:8-16 (2000).
- Hayward RE, Derisi JL, Alfadhi S, Kaslow DC, Brown PO, Rathod PK. Shotgun DNA microarrays and stage-specific gene expression in *Plasmodium falciparum* malaria. *Mol Microbiol* 35:6-14 (2000).
- Iyer VR, Eisen MB, Ross DT, Schuler G, Moore T, Lee JCF, Trent JM, Staudt LM, Hudson J Jr, Boguski MS, et

- al. The transcriptional program in the response of human fibroblasts to serum. *Science* 283:83–87 (1999).
45. Walker J, Riggley K. Gene expression profiling in human peripheral blood mononuclear cells using high-density filter-based cDNA microarrays. *J Immunol Methods* 239:167–179 (2000).
 46. Guo X, Liao K. Analysis of gene expression profile during 3T3-L1 preadipocyte differentiation. *Gene* 251:45–53 (2000).
 47. Funk WD, Wang CK, Shelton DN, Harley CB, Pagon GD, Hoeffler WK. Telomerase expression restores dermal integrity to *in vitro*-aged fibroblasts in a reconstituted skin model. *Exp Cell Res* 258:270–278 (2000).
 48. Greider CW. Telomere length regulation. *Annu Rev Biochem* 65:337–365 (1996).
 49. Blackburn EH. Structure and function of telomeres. *Nature* 350:569–573 (1991).
 50. Harley CB, Futcher AB, Greider CW. Telomeres shorten during ageing of human fibroblasts. *Nature* 345:458–460 (1990).
 51. Lee CK, Klopp RG, Weindruch R, Prolla TA. Gene expression profile of aging and its retardation by caloric restriction. *Science* 285:1390–1393 (1999).
 52. Lashkari DA, DeRisi JL, McCusker JH, Namath AF, Gentile C, Hwang SY, Brown PO, Davis RW. Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc Natl Acad Sci USA* 94:13057–13062 (1997).
 53. Wodicka L, Dong H, Mittmann M, Ho MH, Lockhart DJ. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat Biotechnol* 15:1359–1367 (1997).
 54. Roberts CJ, Nelson B, Marton MJ, Stoughton R, Meyer MR, Bennett HA, He YD, Dai H, Walker WL, Hughes TR, et al. Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. *Science* 287:873–880 (2000).
 55. Press MF. Neuronal development in the cerebellum of lead poisoned neonatal rats. *Acta Neuropathol (Berl)* 40:259–268 (1977).
 56. Markovac J, Goldstein GW. Lead activates protein kinase C in immature rat brain microvessels. *Toxicol Appl Pharmacol* 96:14–23 (1988).
 57. Laterra J, Bressler JP, Indurri RR, Belloni-Olivi L, Goldstein GW. Inhibition of astroglia-induced endothelial differentiation by inorganic lead: a role for protein kinase C. *Proc Natl Acad Sci USA* 89:10748–10752 (1992).
 58. Goldstein GW, Ar D. Lead activates calmodulin sensitive processes. *Life Sci* 33:1001–1006 (1983).
 59. Singh AK. Neurotoxicity in rats chronically exposed to lead ingestion: measurement of intracellular concentrations of free calcium and lead ions in resting or depolarized brain slices. *Neurotoxicology* 16:133–138 (1995).
 60. Hossain MA, Bouton CM, Pevsner J, Laterra J. Induction of vascular endothelial growth factor in human astrocytes by lead. Involvement of a protein kinase C/Activator Protein-1 complex-dependent and hypoxia-inducible factor-1-independent signaling pathway. *J Biol Chem* 275:27874–27882 (2000).
 61. Marton MJ, DeRisi JL, Bennett HA, Iyer VR, Meyer MR, Roberts CJ, Stoughton R, Burchard J, Slade D, Dai H, et al. Drug target validation and identification of secondary drug target effects using DNA microarrays. *Nat Med* 4:1293–1301 (1998).
 62. Alaiya AA, Franzen B, Hagman A, Silfversward C, Moberger B, Linder S, Auer G. Classification of human ovarian tumors using multivariate data analysis of polypeptide expression patterns. *Int J Cancer* 86:731–736 (2000).
 63. Scherf U, Ross DT, Waltham M, Smith LH, Lee JK, Tanabe L, Kohn KW, Reinhold WC, Myers TG, Andrews DT, et al. A gene expression database for the molecular pharmacology of cancer. *Nat Genet* 24:236–244 (2000).
 64. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, Iyer V, Jeffrey SS, Van de Rijn M, Waltham M, et al. Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet* 24:227–235 (2000).
 65. Burczynski ME, McMillian M, Ciervo J, Li L, Parker JB, Dunn RT II, Hicken S, Farr S, Johnson MD. Toxicogenomics-based discrimination of toxic mechanism in HepG2 human hepatoma cells. *Toxicol Sci* 58:399–415 (2000).
 66. Hamadeh HK, Bushel P, Jayadev S, Martin K, DiSorbo O, Sieber S, Bennett L, Tennant R, Stoll R, Barrett JC, et al. Unpublished data.
 67. National Center for Biotechnology Information. Available at: www.ncbi.nlm.nih.gov [accessed 1 November 2000].
 68. Bentley DR. The Human Genome Project—an overview. *Med Res Rev* 20:189–196 (2000).
 69. Afshari CA, Nuwaysir EF, Barrett JC. Application of complementary DNA microarray technology to carcinogen identification, toxicology, and drug safety evaluation. *Cancer Res* 59:4759–4760 (1999).
 70. Cheung VG, Morley M, Aguilar F, Massimi A, Kucherlapati R, Childs G. Making and reading microarrays. *Nat Genet* 21:15–19 (1999).
 71. Ermolaeva O, Rastogi M, Pruitt KD, Schuler GD, Bittner ML, Chen Y, Simon R, Meltzer P, Trent JM, Boguski MS. Data management and analysis for gene expression arrays. *Nat Genet* 20:19–23 (1998).
 72. Chen Y, Dougherty E, Bittner M. Ratio-based decisions and the quantitative analysis of cDNA microarray images. *J Biomed Optics* 2:364–374 (1997).
 73. Wittes J, Friedman HP. Searching for evidence of altered gene expression: a comment on statistical analysis of microarray data. *J Natl Cancer Inst* 91:400–401 (1999).
 74. Lee ML, Kuo FC, Whitmore GA, Sklar J. Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. *Proc Natl Acad Sci USA* 97:9834–9839 (2000).
 75. Casella G, Berger RL. *Statistical Inference*. Pacific Grove, CA:Wadsworth Inc., 1990.
 76. Wolfinger RD. Unpublished data.
 77. Kerr MK, Martin M, Churchill GA. Analysis of variance for gene expression microarray data. *J Comput Biol* 7:819–837 (2000).
 78. Hughes TR, Marton MJ, Jones AR, Roberts CJ, Stoughton R, Armour CD, Bennett HA, Coffey E, Dai H, He YD, et al. Functional discovery via a compendium of expression profiles. *Cell* 102:109–126 (2000).
 79. Tamayo P, Slonim D, Mesirov J, Zhu Q, Kitareewan S, Dmitrovsky E, Lander ES, Golub TR. Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc Natl Acad Sci USA* 96:2907–2912 (1999).
 80. Li L, Darden TA, Weinberg CR, Pedersen LG. Gene assessment and sample classification for gene expression data using a genetic algorithm/k-nearest neighbor method. *Comb Chem High Throughput Screen* (in press).
 81. Barlund M, Forozan F, Kononen J, Bubendorf L, Chen Y, Bittner ML, Torhorst J, Haas P, Bucher C, Sauter G, et al. Detecting activation of ribosomal protein S6 kinase by complementary DNA and tissue microarray analysis. *J Natl Cancer Inst* 92:1252–1259 (2000).
 82. Forozan F, Mahlamaki EH, Monni O, Chen Y, Veldman R, Jiang Y, Gooden GC, Ethier SP, Kallioniemi A, Kallioniemi OP. Comparative genomic hybridization analysis of 38 breast cancer cell lines: a basis for interpreting complementary DNA microarray data. *Cancer Res* 60:4519–4525 (2000).
 83. Kudoh K, Ramanna M, Ravatn R, Elkahlon AG, Bittner ML, Meltzer PS, Trent JM, Dalton WS, Chin KV. Monitoring the expression profiles of doxorubicin-induced and doxorubicin-resistant cancer cells by cDNA microarray. *Cancer Res* 60:4161–4166 (2000).
 84. Lehman TA, Haffty BG, Carbone CJ, Bishop LR, Gumbs AA, Krishnan S, Shields PG, Modali R, Turner BC. Elevated frequency and functional activity of a specific germ-line p53 intron mutation in familial breast cancer. *Cancer Res* 60:1062–1069 (2000).
 85. Lyakhovich A, Aksenov N, Pennanen P, Miettinen S, Ahonen MH, Syvala H, Ylikomi T, Tuohimaa P. Vitamin D induced up-regulation of keratinocyte growth factor (FGF-7/KGF) in MCF-7 human breast cancer cells. *Biochem Biophys Res Commun* 273:675–680 (2000).
 86. Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* 23:41–46 (1999).
 87. Yang GP, Ross DT, Kuang WW, Brown PO, Weigel RJ. Combining SSH and cDNA microarrays for rapid identification of differentially expressed genes. *Nucleic Acids Res* 27:1517–1523 (1999).
 88. Hacia JG, Woski SA, Fidanza J, Edgemon K, Hunt N, McCall G, Fodor SP, Collins FS. Enhanced high density oligonucleotide array-based sequence analysis using modified nucleoside triphosphates. *Nucleic Acids Res* 26:4975–4982 (1998).
 89. Barlund M, Monni O, Kononen J, Cornelson R, Torhorst J, Sauter G, Kallioniemi O-P, Kallioniemi A. Multiple genes at 17q23 undergo amplification and overexpression in breast cancer. *Cancer Res* 60:5340–5344 (2000).
 90. Clement JH, Marr N, Meissner A, Schwalbe M, Sebald W, Kliche KO, Hoffken K, Wolff S. Bone morphogenetic protein 2 (BMP-2) induces sequential changes of Id gene expression in the breast cancer cell line MCF-7. *J Cancer Res Clin Oncol* 126:271–279 (2000).
 91. Favis R, Barany F. Mutation detection in K-ras, BRCA1, BRCA2, and p53 using PCR/LDR and a universal DNA microarray. *Ann NY Acad Sci* 906:39–43 (2000).
 92. Kauraniemi P, Hedenfalk I, Persson K, Duggan DJ, Tanner M, Johannsson O, Olsson H, Trent JM, Isola J, Borg A. MYB oncogene amplification in hereditary BRCA1 breast cancer. *Cancer Res* 60:5323–5328 (2000).
 93. Nacht M, Ferguson AT, Zhang W, Petroziello JM, Cook BP, Gao YH, Maguire S, Riley D, Coppola G, Landes GM, et al. Combining serial analysis of gene expression and array technologies to identify genes differentially expressed in breast cancer. *Cancer Res* 59:5464–5470 (1999).
 94. Pinkel D, Seagraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 20:207–211 (1998).
 95. Perou CM, Jeffrey SS, van de Rijn M, Rees CA, Eisen MB, Ross DT, Pergamenschikov A, Williams CF, Zhu SX, Lee JC, et al. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc Natl Acad Sci USA* 96:9212–9217 (1999).
 96. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, et al. Molecular portraits of human breast tumours. *Nature* 406:747–752 (2000).
 97. Wang K, Gan L, Jeffery E, Gayle M, Gown AM, Skelly M, Nelson PS, Ng WW, Schummer M, Hood L, et al. Monitoring gene expression profile changes in ovarian carcinomas using cDNA microarray. *Gene* 229:101–108 (1999).
 98. Ono K, Tanaka T, Tsunoda T, Kitahara O, Kihara C, Okamoto A, Ochiai K, Takagi T, Nakamura Y. Identification by cDNA microarray of genes involved in ovarian carcinogenesis. *Cancer Res* 60:5007–5011 (2000).
 99. Shim C, Zhang W, Rhee CH, Lee JH. Profiling of differentially expressed genes in human primary cervical cancer by complementary DNA expression array. *Clin Cancer Res* 4:3045–3050 (1998).
 100. Smid-Koopman E, Blok LJ, Chadha-Ajwani S, Helmerhorst TJ, Brinkmann AO, Huikeshoven FJ. Gene expression profiles of human endometrial cancer samples using a cDNA-expression array technique: assessment of an analysis method. *Br J Cancer* 83:246–251 (2000).
 101. Elek J, Park KH, Narayanan R. Microarray-based expression profiling in prostate tumors. *In Vivo* 14:173–182 (2000).
 102. Perrone EE, Theoharis C, Mucci NR, Hayasaka S, Taylor JM, Cooney KA, Rubin MA. Tissue microarray assessment of prostate cancer tumor proliferation in African-American and white men. *J Natl Cancer Inst* 92:937–939 (2000).
 103. Svaren J, Ehrig T, Abdulkadir SA, Ehrenguber MU, Watson MA, Milbrandt J. EGR1 Target genes in prostate carcinoma cells identified by microarray analysis. *J Biol Chem* 275(49):38524–38531 (2000).
 104. Vaarala MH, Porvari K, Kyllonen A, Vihko P. Differentially expressed genes in two LNCaP prostate cancer cell lines reflecting changes during prostate cancer progression. *Lab Invest* 80:1259–1268 (2000).
 105. Xu J, Stolk JA, Zhang X, Silva SJ, Houghton RL, Matsumura M, Vedvick TS, Leslie KB, Badaro R, Reed SG. Identification of differentially expressed genes in human prostate cancer using subtraction and microarray. *Cancer Res* 60:1677–1682 (2000).
 106. Bubendorf L, Kolmer M, Kononen J, Koivisto P, Mousse S, Chen Y, Mahlamaki E, Schraml P, Moch H, Willi N, et al. Hormone therapy failure in human prostate cancer: analysis by complementary DNA and tissue microarrays. *J Natl Cancer Inst* 91:1758–1764 (1999).
 107. Bubendorf L, Kononen J, Koivisto P, Schraml P, Moch H, Gasser TC, Willi N, Mihatsch MJ, Sauter G, Kallioniemi OP. Survey of gene amplifications during prostate cancer progression by high-throughput fluorescence in situ hybridization on tissue microarrays. *Cancer Res* 59:803–806 (1999).
 108. Carlisle AJ, Prabhu VV, Elkahlon A, Hudson J, Trent JM, Linehan WM, Williams ED, Emmert-Buck MR, Liotta LA, Munson PJ, et al. Development of a prostate cDNA microarray and statistical gene expression analysis package. *Mol Carcinog* 28:12–22 (2000).

109. Nelson PS, Han D, Rochon Y, Corthals GL, Lin B, Monson A, Nguyen V, Franza BR, Plymate SR, Aebbersold R, et al. Comprehensive analyses of prostate gene expression: convergence of expressed sequence tag databases, transcript profiling and proteomics. *Electrophoresis* 21:1823–1831 (2000).
110. Mariadason JM, Corner GA, Augenlicht LH. Genetic reprogramming in pathways of colonic cell maturation induced by short chain fatty acids: comparison with trichostatin A, sulindac, and curcumin and implications for chemoprevention of colon cancer. *Cancer Res* 60:4561–4572 (2000).
111. Rechreche H, Montalto G, Mallo GV, Vasseur S, Marasa L, Soubeyran P, Dagorn JC, Iovanna JL. Pap, reg I alpha and reg I beta mRNAs are concomitantly up-regulated during human colorectal carcinogenesis. *Int J Cancer* 81:688–694 (1999).
112. Sturniolo T, Bono E, Ding J, Raddizzani L, Tuereci O, Sahin U, Braxenthaler M, Gallazzi F, Protti MP, Sinigaglia F, et al. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nat Biotechnol* 17:555–561 (1999).
113. Mallo GV, Rechreche H, Frigerio JM, Rocha D, Zweibaum A, Lacasa M, Jordan BR, Duseti NJ, Dagorn JC, Iovanna JL. Molecular cloning, sequencing and expression of the mRNA encoding human Cdx1 and Cdx2 homeobox. Down-regulation of Cdx1 and Cdx2 mRNA expression during colorectal carcinogenesis. *Int J Cancer* 74:35–44 (1997).
114. Alon U, Barkai N, Notterman DA, Gish K, Ybarra S, Mack D, Levine AJ. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proc Natl Acad Sci USA* 96:6745–6750 (1999).
115. Clarke PA, Hostein I, Banerji U, Stefano FD, Maloney A, Walton M, Judson I, Workman P. Gene expression profiling of human colon cancer cells following inhibition of signal transduction by 17-allylamino-17-demethoxygeldanamycin, an inhibitor of the hsp90 molecular chaperone. *Oncogene* 19:4125–4133 (2000).
116. Geng MM, Ellenrieder V, Wallrapp C, Muller-Pillasch F, Sommer G, Adler G, Gress TM. Use of representational difference analysis to study the effect of TGF β on the expression profile of a pancreatic cancer cell line. *Genes Chromosomes Cancer* 26:70–79 (1999).
117. Anderson KM, Alrefai WA, Bonomi P, Dudeja P, Ou D, Anderson C, Harris JE. Altered oncogene, tumor suppressor and cell-cycle gene expression in PANC-1 cells cultured with the pleiotropic 5-lipoxygenase inhibitor, MK886, assessed with a gene chip. *Anticancer Res* 19:3873–3887 (1999).
118. Moch H, Schraml P, Bubendorf L, Mirlacher M, Kononen J, Gasser T, Mihatsch MJ, Kallioniemi OP, Sauter G. High-throughput tissue microarray analysis to evaluate genes uncovered by cDNA microarray screening in renal cell carcinoma. *Am J Pathol* 154:981–986 (1999).
119. Lam LT, Ronchini C, Norton J, Capobianco AJ, Bresnick EH. Suppression of erythroid but not megakaryocytic differentiation of human K562 erythroleukemic cells by notch-1. *J Biol Chem* 275:19676–19684 (2000).
120. Rusiniak ME, Yu M, Ross DT, Tolhurst EC, Slack JL. Identification of B94 (TNFAIP2) as a potential retinoic acid target gene in acute promyelocytic leukemia. *Cancer Res* 60:1824–1829 (2000).
121. Cossman J, Annunziata CM, Barash S, Staudt L, Dillon P, He WW, Ricciardi-Castagnoli P, Rosen CA, Carter KC. Reed-Sternberg cell gene expression supports a B-cell lineage. *Blood* 94:411–416 (1999).
122. Kapp U, Yeh WC, Patterson B, Elia AJ, Kagi D, Ho A, Hessel A, Tipsword M, Williams A, Mirtsos C, et al. Interleukin 13 is secreted by and stimulates the growth of Hodgkin and Reed-Sternberg cells. *J Exp Med* 189:1939–1946 (1999).
123. Amundson SA, Bittner M, Chen Y, Trent J, Meltzer P, Fornace AJ Jr. Fluorescent cDNA microarray hybridization reveals complexity and heterogeneity of cellular genotoxic stress responses. *Oncogene* 18:3666–3672 (1999).
124. Amundson SA, Do KT, Shahab S, Bittner M, Meltzer P, Trent J, Fornace AJ Jr. Identification of potential mRNA biomarkers in peripheral blood lymphocytes for human exposure to ionizing radiation. *Radiat Res* 154:342–346 (2000).
125. Bishay K, Ory K, Lebeau J, Levalois C, Olivier MF, Chevillard S. DNA damage-related gene expression as biomarkers to assess cellular response after gamma irradiation of a human lymphoblastoid cell line. *Oncogene* 19:916–923 (2000).
126. Huang P, Feng L, Oldham EA, Keating MJ, Plunkett W. Superoxide dismutase as a target for the selective killing of cancer cells. *Nature* 407:390–395 (2000).
127. Loberg LI, Engdahl WR, Gauger JR, McCormick DL. Expression of cancer-related genes in human cells exposed to 60 Hz magnetic fields. *Radiat Res* 153:679–684 (2000).
128. Lossos IS, Alizadeh AA, Eisen MB, Chan WC, Brown PO, Botstein D, Staudt LM, Levy R. Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. *Proc Natl Acad Sci USA* 97:10209–10213 (2000).
129. Shaffer AL, Yu X, He Y, Boldrick J, Chan EP, Staudt LM. BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control. *Immunity* 13:199–212 (2000).
130. Tsou AP, Wu KM, Tsen TY, Chi CW, Chiu JH, Lui WY, Hu CP, Chang C, Chou CK, Tsai SF. Parallel hybridization analysis of multiple protein kinase genes: identification of gene expression patterns characteristic of human hepatocellular carcinoma. *Genomics* 50:331–340 (1998).
131. Akama TO, Okazaki Y, Ito M, Okuizumi H, Konno H, Muramatsu M, Plass C, Held WA, Hayashizaki Y. Restriction landmark genomic scanning (RLGS-M)-based genome-wide scanning of mouse liver tumors for alterations in DNA methylation status. *Cancer Res* 57:3294–3299 (1997).
132. Richter J, Wagner U, Kononen J, Fijan A, Bruderer J, Schmid U, Ackermann D, Maurer R, Alund G, Kononagel H, et al. High-throughput tissue microarray analysis of cyclin E gene amplification and overexpression in urinary bladder cancer. *Am J Pathol* 157:787–794 (2000).
133. Wikman FP, Lu ML, Thykjaer T, Olesen SH, Andersen LD, Cordon-Cardo C, Orntoft TF. Evaluation of the performance of a p53 sequencing microarray chip using 140 previously sequenced bladder tumor samples. *Clin Chem* 46:1555–1561 (2000).
134. Villaret DB, Wang T, Dillon D, Xu J, Sivam D, Cheever MA, Reed SG. Identification of genes overexpressed in head and neck squamous cell carcinoma using a combination of complementary DNA subtraction and microarray analysis. *Laryngoscope* 110:374–381 (2000).
135. Ohyama H, Zhang X, Kohno Y, Alevizos I, Posner M, Wong DT, Todd R. Laser capture microdissection-generated target sample for high-density oligonucleotide array hybridization. *Biotechniques* 29:530–536 (2000).
136. Leethanakul C, Patel V, Gillespie J, Pallente M, Enslay JF, Koontongkaew S, Liotta LA, Emmert-Buck M, Gutkind JS. Distinct pattern of expression of differentiation and growth-related genes in squamous cell carcinomas of the head and neck revealed by the use of laser capture microdissection and cDNA arrays. *Oncogene* 19:3220–3224 (2000).
137. Xie L, Xu L, He Z, Zhou W, Wang L, Zhang L, Lan K, Ren C, Liu W, Yao K. Identification of differentially expressed genes in nasopharyngeal carcinoma by means of the Atlas human cancer cDNA expression array. *J Cancer Res Clin Oncol* 126:400–406 (2000).
138. Liau LM, Lallone RL, Seitz RS, Buznikov A, Gregg JP, Kornblum HL, Nelson SF, Bronstein JM. Identification of a human glioma-associated growth factor gene, granulin, using differential immunoprecipitation. *Cancer Res* 60:1353–1360 (2000).
139. Lal A, Lash AE, Altschul SF, Velculescu V, Zhang L, McLendon RE, Marra MA, Prange C, Morin PJ, Polyak K et al. A public database for gene expression in human cancers. *Cancer Res* 59:5403–5407 (1999).
140. Takahashi N, Hashida H, Zhao N, Misumi Y, Sakaki Y. High-density cDNA filter analysis of the expression profiles of the genes preferentially expressed in human brain. *Gene* 164:219–227 (1995).
141. Fuller GN, Rhee CH, Hess KR, Caskey LS, Wang R, Bruner JM, Yung WK, Zhang W. Reactivation of insulin-like growth factor binding protein 2 expression in glioblastoma multiforme: a revelation by parallel gene expression profiling. *Cancer Res* 59:4228–4232 (1999).
142. Rhee CH, Hess K, Jabbur J, Ruiz M, Yang Y, Chen S, Chenchik A, Fuller GN, Zhang W. cDNA expression array reveals heterogeneous gene expression profiles in three glioblastoma cell lines. *Oncogene* 18:2711–2717 (1999).
143. Loging WT, Lal A, Siu IM, Loney TL, Wikstrand CJ, Marra MA, Prange C, Bigner DD, Strausberg RL, Riggins GJ. Identifying potential tumor markers and antigens by database mining and rapid expression screening. *Genome Res* 10:1393–1402 (2000).
144. Hong TM, Yang PC, Peck K, Chen JJ, Yang SC, Chen YC, Wu CW. Profiling the downstream genes of tumor suppressor PTEN in lung cancer cells by complementary DNA microarray. *Am J Respir Cell Mol Biol* 23:355–363 (2000).
145. Wang T, Hopkins D, Schmidt C, Silva S, Houghton R, Takita H, Repasky E, Reed SG. Identification of genes differentially over-expressed in lung squamous cell carcinoma using combination of cDNA subtraction and microarray analysis. *Oncogene* 19:1519–1528 (2000).
146. Ahrendt SA, Chow JT, Yang SC, Wu L, Zhang MJ, Jen J, Sidransky D. Alcohol consumption and cigarette smoking increase the frequency of p53 mutations in non-small cell lung cancer. *Cancer Res* 60:3155–3159 (2000).
147. Ahrendt SA, Halachmi S, Chow JT, Wu L, Halachmi N, Yang SC, Wehage S, Jen J, Sidransky D. Rapid p53 sequence analysis in primary lung cancer using an oligonucleotide probe array. *Proc Natl Acad Sci USA* 96:7382–7387 (1999).
148. Sato H, Sagai M, Suzuki KT, Aoki Y. Identification, by cDNA microarray, of A-raf and proliferating cell nuclear antigen as genes induced in rat lung by exposure to diesel exhaust. *Res Commun Mol Pathol Pharmacol* 105:77–86 (1999).
149. Sandhu H, Dehnen W, Roller M, Abel J, Unfried K. mRNA expression patterns in different stages of asbestos-induced carcinogenesis in rats. *Carcinogenesis* 21:1023–1029 (2000).
150. Su YA, Bittner ML, Chen Y, Tao L, Jiang Y, Zhang Y, Stephan DA, Trent JM. Identification of tumor-suppressor genes using human melanoma cell lines UACC903, UACC903(+6), and SRS3 by comparison of expression profiles. *Mol Carcinog* 28:119–127 (2000).
151. Maniatis AJ, Folberg R, Hess A, Seftor EA, Gardner LM, Pe'er J, Trent JM, Meltzer PS, Hendrix MJ. Vascular channel formation by human melanoma cells in vivo and in vitro: vasculogenic mimicry. *Am J Pathol* 155:739–752 (1999).
152. Clark EA, Golub TR, Lander ES, Hynes RO. Genomic analysis of metastasis reveals an essential role for RhoC. *Nature* 406:532–535 (2000).
153. Khan J, Simon R, Bittner M, Chen Y, Leighton SB, Pohida T, Smith PD, Jiang Y, Gooden GC, Trent JM, et al. Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. *Cancer Res* 58:5009–5013 (1998).
154. Khan J, Bittner ML, Saal LH, Teichmann U, Azorsa DO, Gooden GC, Pavan WJ, Trent JM, Meltzer PS. cDNA microarrays detect activation of a myogenic transcription program by the PAX3-FKHR fusion oncogene. *Proc Natl Acad Sci USA* 96:13264–13269 (1999).
155. Pandita A, Zielenska M, Thorne P, Bayani J, Godbout R, Greenberg M, Squire JA. Application of comparative genomic hybridization, spectral karyotyping, and microarray analysis in the identification of subtype-specific patterns of genomic changes in rhabdomyosarcoma. *Neoplasia* 1:262–275 (1999).
156. Scholl FA, Betts DR, Niggli FK, Schafer BW. Molecular features of a human rhabdomyosarcoma cell line with spontaneous metastatic progression. *Br J Cancer* 82:1239–1245 (2000).
157. Welford SM, Gregg J, Chen E, Garrison D, Sorensen PH, Denny CT, Nelson SF. Detection of differentially expressed genes in primary tumor tissues using representational differences analysis coupled to microarray hybridization. *Nucleic Acids Res* 26:3059–3065 (1998).
158. Schraml P, Kononen J, Bubendorf L, Moch H, Bissig H, Nocito A, Mihatsch MJ, Kallioniemi OP, Sauter G. Tissue microarrays for gene amplification surveys in many different tumor types. *Clin Cancer Res* 5:1966–1975 (1999).
159. Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G, Kallioniemi OP. Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* 4:844–847 (1998).
160. Solinas-Toldo S, Lampel S, Stilgenbauer S, Nicklenko J, Benner A, Dohner H, Cremer T, Lichter P. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* 20:399–407 (1997).
161. Gerry NP, Witowski NE, Day J, Hammer RP, Barany G, Barany F. Universal DNA microarray method for multiplex detection of low abundance point mutations. *J Mol Biol* 292:251–262 (1999).

162. Belcher CE, Drenkow J, Kehoe B, Gingeras TR, McNamara N, Lemjabbar H, Basbaum C, Relman DA. The transcriptional responses of respiratory epithelial cells to *Bordetella pertussis* reveal host defensive and pathogen counter-defensive strategies. *Proc Natl Acad Sci USA* 97:13847–13852 (2000).
163. Holden PR, James NH, Brooks AN, Roberts RA, Kimber I, Pennie WD. Identification of a possible association between carbon tetrachloride-induced hepatotoxicity and interleukin-8 expression. *J Biochem Mol Toxicol* 14:283–290 (2000).
164. Bartosiewicz M, Trounstein M, Barker D, Johnston R, Buckpitt A. Development of a toxicological gene array and quantitative assessment of this technology. *Arch Biochem Biophys* 376:66–73 (2000).
165. Puga A, Maier A, Medvedovic M. The transcriptional signature of dioxin in human hepatoma HepG2 cells. *Biochem Pharmacol* 60:1129–1142 (2000).
166. Gross C, Kelleher M, Iyer VR, Brown PO, Winge DR. Identification of the copper regulon in *Saccharomyces cerevisiae* by DNA Microarrays. *J Biol Chem* 275:32310–32316 (2000).
167. DeRisi J, van den Hazel B, Marc P, Balzi E, Brown P, Jacq C, Goffeau A. Genome microarray analysis of transcriptional activation in multidrug resistance yeast mutants. *FEBS Lett* 470:156–160 (2000).
168. Wilson M, DeRisi J, Kristensen HH, Imboden P, Rane S, Brown PO, Schoolnik GK. Exploring drug-induced alterations in gene expression in *Mycobacterium tuberculosis* by microarray hybridization. *Proc Natl Acad Sci USA* 96:12833–12838 (1999).
169. Cho RJ, Campbell MJ, Winzler EA, Steinmetz L, Conway A, Wodicka L, Wolfsberg TG, Gabrielian AE, Landsman D, Lockhart DJ, et al. A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol Cell* 2:65–73 (1998).
170. Posas F, Chambers JR, Heyman JA, Hoeffler JP, de Nadal E, Arino J. The transcriptional response of yeast to saline stress. *J Biol Chem* 275:17249–17255 (2000).
171. Ferea TL, Brown PO. Observing the living genome. *Curr Opin Genet Dev* 9:715–722 (1999).
172. Bonnet C, Perret E, Dumont X, Picard A, Caput D, Lenaers G. Identification and transcription control of fission yeast genes repressed by an ammonium starvation growth arrest. *Yeast* 16:23–33 (2000).
173. Lee TI, Causton HC, Holstege FC, Shen WC, Hannett N, Jennings EG, Winston F, Green MR, Young RA. Redundant roles for the TFIID and SAGA complexes in global transcription. *Nature* 405:701–704 (2000).
174. Hughes TR, Roberts CJ, Dai H, Jones AR, Meyer MR, Slade D, Burchard J, Dow S, Ward TR, Kidd MJ, et al. Widespread aneuploidy revealed by DNA microarray expression profiling. *Nat Genet* 25:333–337 (2000).
175. Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA. Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* 95:717–728 (1998).
176. Galitski T, Saldanha AJ, Styles CA, Lander ES, Fink GR. Ploidy regulation of gene expression. *Science* 285:251–254 (1999).
177. Lyons TJ, Gasch AP, Gaither LA, Botstein D, Brown PO, Eide DJ. Genome-wide characterization of the Zap1p zinc-responsive regulon in yeast. *Proc Natl Acad Sci USA* 97:7957–7962 (2000).
178. O'Brien T, Tjian R. Different functional domains of TAFII250 modulate expression of distinct subsets of mammalian genes. *Proc Natl Acad Sci USA* 97:2456–2461 (2000).
179. Sudarsanam P, Iyer VR, Brown PO, Winston F. Whole-genome expression analysis of *snf/swi* mutants of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci USA* 97:3364–3369 (2000).
180. Robertson LS, Causton HC, Young RA, Fink GR. The yeast A kinases differentially regulate iron uptake and respiratory function. *Proc Natl Acad Sci USA* 97:5984–5988 (2000).
181. Wyrick JJ, Holstege FC, Jennings EG, Causton HC, Shore D, Grunstein M, Lander ES, Young RA. Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. *Nature* 402:418–421 (1999).
182. Gray NS, Wodicka L, Thunnissen AM, Norman TC, Kwon S, Espinoza FH, Morgan DO, Barnes G, LeClerc S, Meijer L, et al. Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science* 281:533–538 (1998).
183. Pendurthi UR, Allen KE, Ezban M, Rao LV. Factor VIIa and thrombin induce the expression of Cyr61 and connective tissue growth factor, extracellular matrix signaling proteins that could act as possible downstream mediators in factor VIIa x tissue factor-induced signal transduction. *J Biol Chem* 275:14632–14641 (2000).
184. Kelly DL, Rizzino A. DNA microarray analyses of genes regulated during the differentiation of embryonic stem cells. *Mol Reprod Dev* 56:113–123 (2000).
185. Dietz AB, Bulur PA, Knutson GJ, Maticic R, Vuk-Pavlovic S. Maturation of human monocyte-derived dendritic cells studied by microarray hybridization. *Biochem Biophys Res Commun* 275:731–738 (2000).
186. Webb GC, Akbar MS, Zhao C, Steiner DF. Expression profiling of pancreatic beta cells: glucose regulation of secretory and metabolic pathway genes. *Proc Natl Acad Sci USA* 97:5773–5778 (2000).
187. Feng Y, Yang JH, Huang H, Kennedy SP, Turi TG, Thompson JF, Libby P, Lee RT. Transcriptional profile of mechanically induced genes in human vascular smooth muscle cells. *Circ Res* 85:1118–1123 (1999).
188. Tanaka TS, Jaradat SA, Lim MK, Kargul GJ, Wang X, Grahovac MJ, Pantano S, Sano Y, Piao Y, Nagaraja R, et al. Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. *Proc Natl Acad Sci USA* 97:9127–9132 (2000).
189. White KP, Rifkin SA, Hurban P, Hogness DS. Microarray analysis of *Drosophila* development during metamorphosis. *Science* 286:2179–2184 (1999).
190. Rogge L, Bianchi E, Biffi M, Bono E, Chang SY, Alexander H, Santini C, Ferrari G, Sinigaglia L, Seiler M, et al. Transcript imaging of the development of human T helper cells using oligonucleotide arrays. *Nat Genet* 25:96–101 (2000).
191. Phillips RL, Ernst RE, Brunk B, Ivanova N, Mahan MA, Deanehan JK, Moore KA, Overton GC, Lemischka IR. The genetic program of hematopoietic stem cells. *Science* 288:1635–1640 (2000).